

# **ESIP Federation**

## **Insights on Technology**

Rob Raskin

NASA/Jet Propulsion Laboratory  
California Institute of Technology

# Outline

- Part 1: Federation context for technology development
- Part 2: A tour of technologies developed by ESIP members
- Part 3: SEEDS-related activities within the Federation

# 1. Federation Context



# Historical Context of Federation

- Federation created to be a self-governing entity
  - Governed by constitution, bylaws, elected officers
  - Define its own direction
  - Provide an alternative to rigidity of ECS
- One of the first resolutions was:
  - “No requirement shall be imposed on an ESIP without its consent”

# NASA-Imposed Requirements

- ESIPs were required to submit DIF entries to GCMD for all their products and services
  - This satisfied the FGDC metadata requirement for recipients of gov't funds
- A System-Wide Interoperability Layer (SWIL) was to be created by the Federation to enable the member holdings to appear as a unified whole
  - Proposers were to describe in their proposals how this might be carried out

# Creation of the SWIL

- A Federation Interoperability Group (FIG) was formed – it ultimately selected a catalog interoperability system
- SWIL based on DIFs, ESIP Web pages, EDG references
- Mercury and GCMD were the developers
- No additional work required by ESIPs to comply – submittal of DIFs exposed data to SWIL
- Data interoperability was deemed too difficult at the time

# SWIL Development

- FIG became the Interoperability Standing Committee
- SWIL was renamed FIND
- Data interoperability subsequently explored via Clusters
- ESIPs added later were not required to comply

# Interoperability Standing Committee

- Only one of the 9 Federation Committees that has technology as its primary focus
- Organized Technical Workshop at Spring Federation meeting
  - Will have another at Fall meeting (Pasadena)



# Interoperability Standing Committee *(cont.)*

- Exploration of becoming a Technology Committee, to serve larger community needs
- New efforts to look at service and semantic interoperability
- How can we answer the question: Why has data interoperability been so difficult?

# Technology-Oriented Clusters

- DODS Cluster
  - Increased participation in DODS
- GIS Services Cluster (*was Digital Earth*)
  - Assisted ESIPs in installing Web Mapping Servers (WMS)
  - Developed “Guide to Making your Dataset WMS-compliant”
- Content-Based Search Cluster (*no longer active*)
  - Shared expertise in data mining and content-based search
- NewDISS Cluster (*no longer active*)
  - Evolved into Strategic Evolution Working Group



## **2. A Tour of Federation Technologies**

*Based on the ESIP Federation Technical Workshop*

*May 14-15, 2002*

*<http://oceanesip.jpl.nasa.gov/workshop.html>*

# **Federation Technologies:**



## **Data Access**

# DODS

- Server-side read/subset for most data formats
- Client-side integration with most visualization/analysis tools (IDL, MATLAB, VisAD, GrADS)
- About 300 datasets available
- Data Access Protocol (DAP) to be separately developed and distributed
- One of the few ESIPs with a specific mission to work with other ESIPs

# DODS (cont.)

- Advantages

- Integration with science visualization software

- Disadvantages:

- Catalog system remains weak
- Data must be converted to intermediate format for transfer
- User interacts with array row/column parameters rather than geographic parameters

# WMS/WCS

- Open standards developed by Open GIS Consortium (OGC)
  - Web Mapping Server (WMS) for maps;
  - Web Coverage Server (WCS) for data
  - NASA plays major role in standards development processes for WMS/WCS
- Eight WMS or WCS servers in place
  - Advanced by Digital Earth Cluster (*now GIS Services Cluster*)

# WMS/WCS (cont.)

## ● Advantages

- Part of larger suite of standards, e.g. Web Feature Server (WFS) for vector data
- Enables overlay of disparate datasets
- Standards developed in conjunction with broader communities

## ● Disadvantages

- WCS still in development
- Complex data types generally not supported



# MapServer

- Lightweight, public domain GIS
  - alternative to ArcIMS
  - Over 1000 downloads to date
  - Runs on most Unix environments
  - MapScript scripting language
  - Developed at U. Minnesota (TerraSIP)
- Limited GIS functionality
  - Does not enable users to seamlessly link with ESRI software

# **Federation Technologies:**



## **Data Description**

# **Earth Science Markup Language (ESML) (UAH)**

- Specification for XML descriptions of Earth science datasets
- Associated tools to generate XML descriptions
- Associated library to read the data
- Competes with XDF as a description language

# **Federation Technologies:**



## **Data Management**

# Earth System Science Workbench (ESSW)

- Based loosely on Sequoia 2000 Project
- Provides recording of parameters
- Provides client “notebook” view of end-to-end process
- Currently used only at UCSB

# **BigSur (ScienceTools Corp.)**

- Also has roots in Sequoia 2000
- Database-centric approach
  - Put programs, parameters, and data in DBMS
- Pure Java
- Currently used at Langley DAAC
  - Also tested at Ocean ESIP 1997-2000

# **Federation Technologies:**



## **Data/Knowledge Discovery**

# Federation Interactive Network for Discovery (FIND)

- GCMD developed portal to Federation holdings, accessible both from Federation and GCMD pages
- Mercury developed portal to:
  - Federation GCMD listings
  - Federation Web pages



# THE FEDERATION OF EARTH SCIENCE INFORMATION PARTNERS

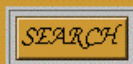
HOME KNOWLEDGE CENTER EDUCATION CENTER TECHNOLOGY CENTER DATA CENTER

- ▶ ABOUT US
- ▶ VISION STATEMENT
- ▶ NEWS
- ▶ SEARCH FOR DATA
- ▶ PARTNER PROFILES
- ▶ PARTNER WEBSITES
- ▶ CALENDAR
- ▶ GALLERY
- ▶ DIRECTORY
- ▶ CONTACT US
- ▶ FEDERATION BUSINESS
- ▶ SITE MAP

## Search for Federation Data and Services

Use the Federation's Databases and Search Tools to Find Earth Science Data and Services

The Federation maintains a comprehensive inventory of information about its data holdings and services. With over 2,000 data sets available locating those that meet your needs can be a formidable task. You should begin by using the search systems listed below to assist your search. In addition to locating data you can also learn about tools and services that will help you use the data.



◆ Datasets and web pages (Mercury)

◆ Datasets (GCMD)

### Advanced Search

Use Advanced Search to search by space, time, or boolean criteria. Select either the Global Change Master Directory or Mercury Search Engines:

Global Change Master Directory

Mercury

Search:

Search:

- [Data](#)
- [Data by Topic](#)
- [Software & Services](#)

- [Data](#)
- [Web Pages](#)

The Global Change Master Directory is a comprehensive directory of Earth Science and global change data. From the Federation search interfaces, you can explore information for more than 1600 Federation data sets, or access the entire GCMD catalog of data and services.

[more>>](#)



Mercury provides access to information for all ESIP Federation data sets and Web pages through two complementary search mechanisms. Mercury incorporates full GCMD directory entries, Federation Web pages, and additional metadata provided directly to Mercury by the ESIPs.

[more>>](#)

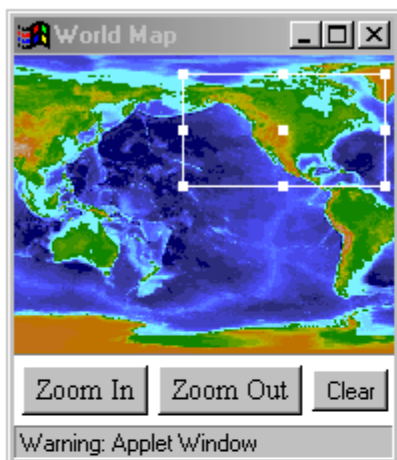
[The EOS Data Gateway](#)  
[Order Data Through the EDG](#)

The EOSDIS Information Management System is a full service, distributed data search and order system. From the EOS Data Gateway, a user can access both high-level and detailed documentation, search through a data product inventory, preview images, and order data. In many instances the data can be directly downloaded to your workstation. EDG also has information about tools for processing, subsetting and viewing the data once you receive it. It is the method of choice for ordering data from NASA's newest Earth Science Enterprise mission - Terra. All ESIP-1s, as well as some ESIP-2s and 3s, participate in this system.



# Global Change Master Directory

[ESIP Free-Text Search](#)



Data sets available through this search are a subset of the Global Change Master Directory and contain *only* those data sets that are part of the the Earth Science Information Partners' (ESIP) Federation.

## Keywords

This field is required.

in

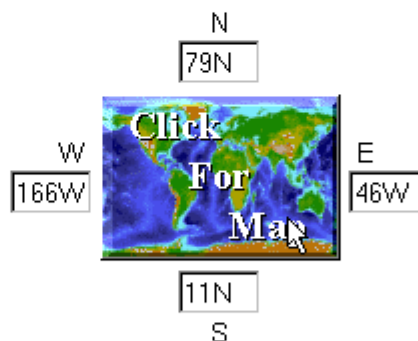
☒ and ☐ or

in

## Spatial Extent

This field is optional.

Include? ☐ YES ☒ NO



## Time Range

This field is optional.

Include? ☐ YES ☒ NO

Time period of interest is  
during

Jan.  1950

through

Feb.  2002

Select the number of records to view:

Select the viewing format:

[Instructions](#)

# FIND Usage

- Mercury: 400 searches/month
- GCMD: 1000 searches/month (includes both Federation and GCMD entry pages)
- EOS article submitted to advertise this search service

# ADaM Data Mining (UAH)

- Suite of tools to carry out data mining in space and time
  - Clustering, pattern recognition, image analysis, filtering, genetic algorithms, selection, texture operations, histograms
- Contributed scientific algorithms become part of system (e.g. cyclone detection)
- Extensive data readers, preprocessing, postprocessing
- Mature technology, but underutilized

# Peer-to-Peer (MODster)

- NAPSTER-like functionality for MODIS data
- Essentially a redirection service enabling users to find MODIS granules of interest
- Appropriate model for cases where multiple sites have similar data product

# THREDDS (Unidata)

- Thematic Realtime Environmental Data Distributed System (THREDDS)
  - Unidata is newest Federation member
- Funded as a DLESE Collections Center
  - Discovery through DLESE discovery tools
- Will provides common catalog and access to datasets accessible through DODS, ADDE, WMS, others
  - Fills in catalog deficiencies of DODS and others
  - Links with existing visulaization tools (e.g. Live Access Server, VisAD)



# WSDL/UDDI

- WSDL and UDDI provide *Web service* interoperability
  - Standard way to access Web services
- Explored by IBM ESIP
  - UDDIs currently for business services

# SWEET (JPL)

- Semantic Web for Earth and Environmental Terminology (SWEET)
- Semantic Web provides *semantic* interoperability
- Enables Web pages to contain XML tags that describe semantic meaning of terms



# **3. SEEDS-Related Activities in the Federation**

# Federation-SEEDS Prototypes

- Last year ('01)
  - 3 proposals funded
    - \$400K total (\$300K from Federation)
  - Funding not yet received
- Current year ('02) and ('03)
  - Combined 2 years => approx. \$600K available
  - RFP released this month
- Projects required to be cross-ESIP

# Technologies in Last Year's Winning Proposals

- Universal Interchange Technology for Earth Science Data (UNITE) (UAH, JPL, ORNL)
  - Plug & play based on ESML descriptors
  - ESML, WCS integration into FIND
- Standards Framework in Support of Dynamic Assembly of NewDISS Components (BASIC, IBM, JPL, ORNL, JHU)
  - WSDL/UDDI, WMS/WCS, FIND integration
- MODster (UCSB, DODS)
  - Peer-to-Peer

# ***Strategic Evolution Working Group***

- Formed at Spring Federation Meeting
  - Created jointly by Products & Services and Interoperability Committees
  - Outgrowth of Federation NewDISS Cluster
- Primary focus is on evolution of Federation's own vision
  - With respect to technical issues
- Intent is to work in parallel with SEEDS and to be a point of contact

# ***Strategic Evolution***

## **Working Group**

### **Objectives**

1. Explore/implement promising technologies, useful in the natural evolution of Federation data systems
2. Further develop the technology
3. Engage customers: Implement technologies that would further facilitate customers use of Federation products and services
4. Identify technology gaps in Federation Services
5. Determine methods and/or standards to facilitate evolving collaborations
6. Document the processes and methods used to achieve evolution

# Conclusions

- Working together in the technology arena to support mutual goals has been challenging
  - This has not been a high priority area for the Federation (relative to SWIL and sustainability)
  - Federation-SEEDS prototypes are a notable exception
  - Standards not always looked upon favorably – positive aspects often overlooked
  - Reactions to ECS have made ESIPs weary of imposed standards

# Conclusions *(cont.)*

- Federation technologies need additional showcasing
  - Data interoperability is our strong area
- Technology standards are important
  - Federation members must demonstrate their value for other members to voluntarily take part